

Millennium Cohort Study

MCS4: A guide to the Linked Education Administrative Datasets

Jon Johnson and Efrosini Setakis

User guide to the data (Third Edition)

April 2015



Centre for Longitudinal Studies

Following lives from birth through the adult years
www.cls.ioe.ac.uk

CLS is an ESRC Resource Centre based at the Institute of Education, London



First published in 2015 by the
Centre for Longitudinal Studies
Institute of Education, University of London
20 Bedford Way
London WC1H 0AL
www.cls.ioe.ac.uk
© Centre for Longitudinal Studies

The Centre for Longitudinal Studies (CLS) is an ESRC Resource Centre based at the Institution of Education. It provides support and facilities for those using the three internationally-renowned birth cohort studies: the National Child Development Study (1958), the 1970 British Cohort Study and the Millennium Cohort Study (2000). CLS conducts research using the birth cohort study data, with a special interest in family life and parenting, family economics, youth life course transitions and basic skills. The views expressed in this work are those of the author(s) (amend as necessary) and do not necessarily reflect the views of the Economic and Social Research Council. All errors and omissions remain those of the author(s).

This document is available in alternative formats.
Please contact the Centre for Longitudinal Studies.
tel: +44 (0)20 7612 6875
email: clsfeedback@ioe.ac.uk

Table of Contents

1. Introduction	1
2. Consent to Education Data Linkage	1
<i>Table 1: Consent Outcomes</i>	2
3. Education Data Linkage	2
a. England.....	2
b. Scotland	2
c. Wales	3
4. Linkage outcome	3
<i>Table 2: Linkage Outcome</i>	3
5. Data variables	3
<i>Table 3: Data Variables (England)</i>	4
<i>Table 4: Data Variables (Scotland)</i>	7
<i>Table 5: Data Variables (Wales)</i>	9

1. Introduction

The Millennium Cohort Study (MCS) is the fourth of Britain's world-renowned national longitudinal birth cohort studies.¹ Each follows a large sample of individuals, born over a specific period of time, through the course of their lives. They provide insight into trajectories/mobility over the life course (e.g. in weight, poverty, education) for the population as a whole, and also into how these movements vary across different sub-populations. They are also uniquely placed to allow for an understanding of the relationship between early life events and circumstances and outcomes later on in life.

The MCS provides detailed information on approximately 19,000 children born at the start of the new century and their families, across the United Kingdom. The cohort members were born, in England and Wales, over the 12 month period starting September 2000; in Scotland and Northern Ireland they were born over 13½ months from November 2000. The sample design allowed for disproportionate representation of families living in areas of child poverty, and in areas of England with high ethnic minority populations. Information was first collected from parents, through a home-based survey, when the cohort members were aged nine months. This first survey recorded, amongst other things, the circumstances of pregnancy and birth, as well as those of the early months of life, and the social and economic background of the children's families.

These multidisciplinary baseline data reveal the diversity of starting points from which the 'Children of the New Century' set out. Subsequent surveys have taken place at ages 3, 5, 7, 11 and 14 (ongoing at time of writing). These surveys coincide with important moments in children's lives, including the preschool period (age 3), and the start, middle and end of primary school (ages 5, 7, 11 respectively). From age 3 onwards, measured physical development and objective cognitive assessments have been carried out with children; surveys also include interviews with both parents (where co-resident), and, increasingly since age 7, with the cohort member. At age 7 (11), a postal survey of class teachers in all four countries (England and Wales) took place.

Further details of the data available from the main survey are available from the CLS website and, in particular, the MCS Guide to the Datasets at www.cls.ioe.ac.uk/mcs.

2. Consent to Education Data Linkage

During the fourth (age 7) sweep, the child's parent / carer was asked for consent to link the child's education records to their MCS survey data. Details of the consent form and the wording of the consent are reproduced in the MCS4 Technical Report, Appendices Section A16: consent form 3. Details of the consent rates are available in the MCS4 Guide to the Education Linkage Dataset 2013.

¹The previous birth cohort studies, still ongoing, are the 1946 National Birth Cohort, the 1958 British Birth Cohort (also known as the National Child Development Study) and the 1970 British Cohort Study.

Table 1: Consent Outcomes

Country of interview	Yes	No	No form	Total
England	8448	293	259	9000
Wales	1898	67	30	1995
Scotland	1536	62	51	1649
Northern Ireland	1288	73	38	1399
Total	13169	495	378	14042

3. Education Data Linkage

During the fourth sweep, the parents / carers were asked the name of the school the cohort child was attending and a lookup table was used by the fieldwork agency to confirm the name and location of the school. This produced a dataset which included the URN of the school, the Local Education Authority and the school number used by the Department of Education in England, the Welsh Assembly, the Scottish Parliament and NISRA in Northern Ireland.

a. England

After some cleaning at CLS to correct incorrectly assigned schools (for example, previous school which has since closed and been renamed), CLS prepared a file of 8445 cases to send to the Department for Education in England (this excluded 3 cases where the interview was in England but the school reported was in another country).

The file contained the following information on the cohort child: Name (first and last), Sex, Date of Birth, Postcode of residence at interview, name of school, URN of current school, Local Education Authority of current school, School Number of current school.

The NPD database team matched the data on the following basis.

- a) Direct match on all variables
- b) Direct match of 4 of the 5 variables or 3 of 4 etc
- c) Fuzzy matches on names – usually that the first 3 or 4 characters are the same – again all manually checked.
- d) Any other combinations we think are safe based on the data supplied.

b. Scotland

After some cleaning at CLS to correct incorrectly assigned schools (for example, previous school which has since closed and been renamed), CLS prepared a file of 1536 cases to send to the Education Analytical Services division within the Learning and Justice Directorate of the Scottish Government.

c. Wales

After some cleaning at CLS to correct incorrectly assigned schools (for example, previous school which has since closed and been renamed), CLS prepared a file of 1890 cases to send to the NHS Wales Informatics Centre (NWIS). This excluded 8 cases where the interview was in Wales but the school reported was in another country. They matched the cohort members to NHS number and this was then sent to SAIL as an anonymised number which was then used to link to the Welsh National Pupil Database.

4. Linkage outcome

Table 2 shows the outcome of the linkages carried out.

Table 2: Linkage Outcome

	England	Wales	Scotland	N. Ireland
No of consents	8448	1898	1536	1288
No of cases sent for linkage	8445	1890	1536	n/a
No. of cases successfully linked	7476*	1890 (1735)**	1407	n/a
% of cases successfully linked	88.5%	100% (92%)	92%	n/a

*Note that the number of linked cases has increased by 635 for England, since the First Edition of the data. This is due to revisions carried out by DfE on a number of cases that did not initially have a successful link.

** The number of cases in Wales linked to NHS number was 100%, however, due to nonresponse by schools to the Census, only 1735 were able to provide individual data.

5. Data variables

In England data are obtained from the Department for Education (DfE) in the form of the National Pupil Database (NPD). The NPD is an extremely rich dataset, containing data collected directly from schools, local authorities, and awarding bodies, alongside data from annual school censuses. Further information can be found on <https://www.gov.uk/government/collections/national-pupil-database>

Linking the NPD provides access to Key Stage 1 and 2 (KS1 and KS2), as well as Pupil Level Annual School Census (PLASC) data. KS1 and KS2 tests are undertaken in all state maintained schools in England, at ages 7 and 10 respectively. Further details can be found on <https://www.gov.uk/national-curriculum/key-stage-1-and-2>

Table 3: Data Variables (England)

Table 3 lists the variables obtained from linkage to the National Pupil Database (NPD) or PLASC.

Variable Name	Description	Source
MCSID	MCS Research ID	CLS
MCSFLAG	Consented to Record Linkage or not	CLS
CMNUM	Cohort Member Number	CLS
CMSEX	CM Sex	CLS
COMPRES1	Child present at MCS1	CLS
COMPRES2	Child present at MCS2	CLS
COMPRES3	Child present at MCS3	CLS
COMPRES4	Child present at MCS4	CLS
COMPRES5	Child present at MCS5	CLS
CONSENT_EDUC	Consent to Educational Linkage (MCS4)	CLS
KS1_LEAANON	Anonymised LEA	CLS
KS1_SCHOOLANON	Anonymised School	CLS
KS1_PLASC_LEAANON	Anonymised LEA as assigned from PLASC data	CLS
KS1_PLASC_SCHOOLANON	Anonymised School as assigned from PLASC data	CLS
KS1_LAESTAB	Anonymised LEA+SCHOOL	NPD
KS1_MONTHOFBIRTH	The month it which the pupil was born	NPD
KS1_ACADYR	KS1 Academic Year	NPD
KS1_TOE_CODE	Type of establishment code (see section 2.5.2)	NPD
KS1_NFTYPE	Institution type generated by PT	NPD
KS1_MMSCH	Maintained mainstream school, academy or CTC	NPD

Variable Name	Description	Source
KS1_MMSCH2	Maintained mainstream school (excluding Academies and CTCs)	NPD
KS1_MSCH	Maintained school, academy or CTC (including maintained special schools)	NPD
KS1_MSCH2	Maintained school, inc maintained special schools (exc Academies and CTCs)	NPD
KS1_MOB1	Pupil joined in latest academic year (year 2)	NPD
KS1_MOB2	Pupil joined in previous academic year (year 1)	NPD
KS1_DISC_SPLI	Multi discount code	NPD
KS1_DISC_READ	Multi discount code	NPD
KS1_DISC_WRIT	Multi discount code	NPD
KS1_DISC_MAT	Multi discount code	NPD
KS1_DISC_SCI	Multi discount code	NPD
KS1_SPEAKANDLISTEN	Nat Curric level awarded for speaking and listening (teacher assessment)	NPD
KS1_READING	Nat Curric level awarded for reading (teacher assessment)	NPD
KS1_WRITING	Nat Curric level awarded for writing (teacher assessment)	NPD
KS1_READWRIT	Average Reading and Writing National Curriculum level	NPD
KS1_MATHS	Nat Curric level awarded for maths (teacher assessment)	NPD
KS1_SCIEXPINVEST	Nat Curric level awarded for scientific enquiry (teacher assessment)	NPD
KS1_SCILIFPROT	Nat Curric level awarded for life and living (teacher assessment)	NPD
KS1_SCIMATPROP	Nat Curric level awarded for materials (teacher assessment)	NPD
KS1_SCIPHYSPROC	Nat Curric level awarded for physical processes (teacher assessment)	NPD
KS1_SCIENCE	Nat Curric level awarded for overall science (teacher assessment)	NPD
KS1_READPOINTS	Reading attainment point score.	NPD
KS1_WRITPOINTS	Writing attainment point score.	NPD
KS1_READWRITPOINTS	Average Reading and Writing attainment point score	NPD

Variable Name	Description	Source
KS1_MATPOINTS	Maths attainment point score.	NPD
KS1_SCIPOINTS	Overall Science attainment point score.	NPD
KS1_APS	Average attainment point score	NPD
KS1_LEVLSPLI	Achieved Level 1 or below (low attainment) in KS1 Speaking and Listening.	NPD
KS1_LEVLREAD	Achieved Level 1 or below (low attainment) in KS1 Reading.	NPD
KS1_LEVLWRIT	Achieved Level 1 or below (low attainment) in KS1 Writing.	NPD
KS1_LEVLMAT	Achieved Level 1 or below (low attainment) in KS1 Maths.	NPD
KS1_LEVLSCI	Achieved Level 1 or below (low attainment) in KS1 overall Science.	NPD
KS1_LEVXSPLI	Achieved Level 2 or above (expected level) in KS1 Speaking and Listening.	NPD
KS1_LEVXREAD	Achieved Level 2 or above (expected level) in KS1 Reading.	NPD
KS1_LEVXWRIT	Achieved Level 2 or above (expected level) in KS1 Writing.	NPD
KS1_LEVXMAT	Achieved Level 2 or above (expected level) in KS1 Maths.	NPD
KS1_LEVXSCI	Achieved Level 2 or above (expected level) in KS1 overall Science.	NPD
KS1_LEV2BREAD	Achieved Level 2B or above in KS1 Reading.	NPD
KS1_LEV2BWRIT	Achieved Level 2B or above in KS1 Writing.	NPD
KS1_LEV2BMAT	Achieved Level 2B or above in KS1 Maths.	NPD
KS1_LEVAXSPLI	Achieved Level 3 or above (above expected level) in KS1 Speaking & Listening.	NPD
KS1_LEVAXREAD	Achieved Level 3 or above (above expected level) in KS1 Reading.	NPD
KS1_LEVAXWRIT	Achieved Level 3 or above (above expected level) in KS1 Writing.	NPD
KS1_LEVAXMAT	Achieved Level 3 or above (above expected level) in KS1 Maths.	NPD
KS1_LEVAXSCI	Achieved Level 3 or above (above expected level) in KS1 overall Science.	NPD
KS1_VERSION	Data version	NPD
RECORDSTATUS_SPR08	Record Status	PLASC

Variable Name	Description	Source
ACADEMICYEAR_SPR08	Academic Year	PLASC
CENSUSDATE_SPR08	Census Date	PLASC
CENSUSTERM_SPR08	Census Term	PLASC
ENROLSTATUS_SPR08	Enrolment status of pupil (e.g. current school, or guest pupil)	PLASC
FSMELIGIBLE_SPR08	Whether the pupil is entitled to free school meals	PLASC
MODEOFTRAVEL_SPR08	Usual mode of transport used by the pupil for getting to school	PLASC
TYPEOFCLASS_SPR08	Type of Class	PLASC
SENPROVISION_SPR08	Provision type under the SEN Code of Practice	PLASC
SPR08_2012_IDACI	IDACI 2010 (Decile)	CLS

Table 4: Data Variables (Scotland)

Variable	Description	Source
MCSID	MCS Research ID	CLS
CMNUM00	Cohort Child Number	CLS
MATCHING	Uniquely identifiable status	CLS
DIFFSCHL	Whether different school listed on the pupil census (1=yes)	CLS
POSSATT	Possible attendance of student in half days	AA
REVATT	Actual attendance of student in half days	AA
TEMPEXCL	Temporary exclusions	AA
STUSTAGE	Stage of the student	PC

Variable	Description	Source
ASNIND	Whether additional support needed (1=yes)	PC
ETHNICBG	Ethnic background	PC
FSMREG	Whether registered for free school meals (1=yes)	PC
LEVELENG	Level of English	PC
FTPRP	Percentage of total FTE that are primary teachers	PC
FTSEP	Percentage of total FTE that are secondary teachers	PC
FTSPP	Percentage of total FTE that are special teachers	PC
PPRIMFSM	Percentage primary free school meal entitled	SMS
PSPECFSM	Percentage special free school meal entitled	SMS
PSECOFSM	Percentage secondary free school meal entitled	SMS
ASNPRIMP	Percentage with additional support needs at primary	PC
ASNSPECP	Percentage with additional support needs at special	PC
EALPRIMP	Percentage with English as another language at primary	PC
EALSPECP	Percentage with English as another language at special	PC
MALEP	Percentage of students male	PC
FEMALEP	Percentage of students female	PC
ASNP	Percentage of students with additional support needs	PC
EALP	Percentage of students with English as another language	PC
SECTOR	Which sector - Primary, special or secondary	PC
PETHNIC	Percentage of Ethnic Minorities	PC
LEAANON	LEA / Board anonymised	CLS
SCHOOLANON	School no, anonymised	CLS

Key to source:

AA (Attendance and Absence for year 2008-2009)
 PC (Pupil Census 2008 for year 2008-2009)
 SMS (School Meals Survey)

Table 5: Data Variables (Wales)

Variable	Description	Source
MCSID	MCS Research ID	CLS
CMNUM00	Cohort Child Number	CLS
ALF_STS	ALF status code in education dataset	PC
ACYEAR	Academic year	PC
FSMELIG	Free school meal eligibility	PC
ENSUB	English teacher assessment level	PC
CYSUB	Welsh teacher assessment level	PC
MASUB	Maths teacher assessment level	PC
SCSUB	Science teacher assessment level	PC
CSI	Did pupil achieve level 2 or above in (ENSUB or CYSUB) and MASUB?	PC
LANG	Language teacher assessment level	PC
LEAANON	LEA / Board anonymised	CLS
SCHOOLANON	School no, anonymised	CLS

Key to source:

PC (Pupil Census 2008 for year 2008-2009)

Centre for Longitudinal Studies
Institute of Education
20 Bedford Way
London WC1H 0AL
Tel: 020 7612 6860
Fax: 020 7612 6880
Email: clsfeedback@ioe.ac.uk
Web: www.cls.ioe.ac.uk